



# Blast 2 Sequences results

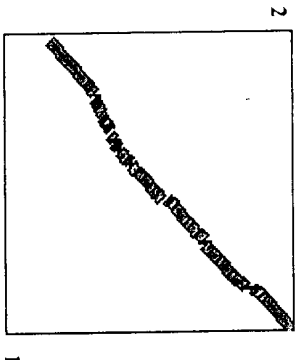
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.1 [Aug-1-2001]

Matrix BLOSUM62 gap open: 11 gap extension: 1  
 x dropout: 5.0 expect: 10.0 wordsize: 3 Filter Align

Sequence 1 gi 728996 COLLAGEN ALPHA 1(XV) CHAIN PRECURSOR Length 1388 (1..1388)

Sequence 2 gi 12644252 COLLAGEN ALPHA 1(XVIII) CHAIN PRECURSOR [CONTAINS: ENDOSTATIN]. Length 1527 (1..1527)



NOTE: The statistics (bitscore and expect value) is calculated based on the size of nr database

Score = 865 bits (2234), Expect = 0.0  
 Identities = 580/1487 (39%), Positives = 716/1487 (48%), Gaps = 348/1487 (23%)

Sequence alignment visualization showing gaps and matches between the two sequences.

```

Query:
      43      LTQLIGDPLPSSVSFVTGYGFPAYSFGPGANVGRPARTLIPSTFFRDFAIRLVKPSST 102
            L QL+GDPLP +S +      PAY FGP +N G+ A+ P FFRDF++ V+P++
      248      LQLIGDPLPEKISQIIDDPHVGPAYIFGPDNSGQVAQYHFPKLFRRDFSLLFHVPRATE 307
            *****
      248      ++++++
      248      *****
      248      *****

Subject:
      103      RGVLFATITDAFOKVIYLGRLSGVEDGHORILLYTTEPGSHVSOEAPAFSVPMTHRMN 162
            GVLFAITDA O V+ LG++LS V DG O I L YTEPG+ +O +F +P +W
      308      AAGVLFAITDAQVAVVSLGVKLSEVRDGOQNISLLYTEPGASQTGTGASFRLPAFVGQWT 367
            *****
      308      *****
      308      *****
      308      *****
      308      *****
      163      RFAMIVGEEVITLVNCEEHSRIIPFORSSQALAESSAGIFMGNAGATGLERFTGSLQOL 222
            *****
    
```

Subject: FA+ V G V L V+CEE R+PF R+SQ L E AG+F+G AG ++F G + +L  
 NONHELIICAL REGION 1 (NC1).  
 COL18A1  
 Mature chain  
 368  
 368 HFAISVDGGSVALYVDCEEFOQVPFARASQGLELRGAGLFVGQAGTADPDKFGMISEL 427  
 368  
 368  
 368  
 Query: 223 TVHDDRTPEELC---DPRESSASGET-SGLQADGVAELLEAVTYTQASPKKAEKVEPI 277  
 V PR C D +E ASG+ SG +E S K K +  
 428 KVRKTPRVSPVHCLEDEDDDEDRAAGDFGSGFEE-----SSKSHKEDTS 471  
 428  
 428  
 428  
 428  
 Subject: NONHELIICAL REGION 1 (NC1).  
 COL18A1  
 Mature chain  
 428  
 428  
 428  
 428  
 Query: 278 NTPPTSSPFEDMELSGEPVPEGTLETTNMSIIQHSKQSGSEILNDTLEGVHSDGDP 337  
 P P P PV +SP G S DP  
 472 LLPGLPQPP-----PV-----TSPPLAG-----STTEDP 496  
 472  
 472  
 472  
 472  
 Subject: NONHELIICAL REGION 1 (NC1).  
 COL18A1  
 Mature chain  
 472  
 472  
 472  
 472  
 Query: 338 ITDSSGAGAFLLDAEKNLAATAAGLAEPVISTAGEAFASSVPTGGPTLSMSTENDEG 397  
 T+ E++ A + G +P + + A ++ +NP G  
 497 RTEE-----TEEDAVIDSIGAETLPGTSSGAMDEAI-----QNPGRG 534  
 497  
 497  
 497  
 497  
 Subject: NONHELIICAL REGION 1 (NC1).  
 COL18A1  
 Mature chain  
 497  
 497  
 497  
 497  
 Query: 398 VTGPDNEERLRATAGAEAFALASMPGEVEASGVAPGELDLMSAQSLEAATVGPSSSED 457  
 + G + GE A G PG G + GP  
 535 LIKG-----GKMGQKGEPGAQG-PPGPA-----GPGPAGPV--- 565  
 535  
 535  
 535  
 535  
 Subject: NONHELIICAL REGION 1 (NC1).  
 COL18A1  
 Mature chain  
 535  
 535  
 535  
 535  
 Query: 458 SLTTAAATEVSLSTEDEEASGVPTDGLPLTATMAPERAIVTSQPGDEEDLAATTEEP 517  
 V + P+ P GP ++  
 566  
 566  
 566  
 566  
 566  
 Subject: NONHELIICAL REGION 2 (NC2).  
 COL18A1  
 Mature chain  
 566  
 566  
 566  
 566  
 Query: 518 LITAGE-ESGSPPPDGPPLPLPTVAPERWITPAQREHVGMKGQAGPKGKGDAGEELPG 576  
 T G + E G P DG P P+ + G G GPKGEKGD G G  
 592 -TPGRDGEPPDGEEDGRP--GDTGQGF-----PGRPGEVGPKEKGDPIGPRG 638  
 592  
 592  
 592  
 592  
 Subject: TRIPLE-HELIICAL REGION 2 (COL2).  
 COL18A1  
 Mature chain  
 592  
 592  
 592  
 592  
 Query: 577 PPEPSGVPVPT-----AGAEAGSGGLGWSGDVSGSGDLVGSSEQLLRGP--PGPPGPP 627  
 PP P GP GP+ + EGSG SGD+ + LRGP PGPPGPP  
 639 PPGPPGPPGPSFRQDKLTIFIDMGSGF-----SGDI-----ESLRGRGPGPPGPP 685

TRIPLE-HELICAL REGION 2 (COL2) .  
COL18A1  
Mature chain  
NONHELICAL REGION 3 (NC3) .  
TRIPLE-HELICAL REGION 3 (COL3)

639 \*\*\*\*\*  
639 ++++++  
639 \*\*\*\*\*  
650 \*\*\*\*\*  
674 \*\*\*\*\*

Query:

Subject:

glycosylation  
TRIPLE-HELICAL REGION 3 (COL3) .  
COL18A1  
Mature chain

628 GLPGIPGKPTDVFMP-----PGSPEDGPAGEPPPPGEGQPGVDGATGLPGMKG 679  
686 G+PG+PG+PG G PG PG+GP G PGP PG G PG +G PG+ G  
686 GVPGLPGEPGRFNGSVAPAGLPGVPGKEGPPGPPPPPPGPPGKEGP--PGVAG 742  
700 \*  
686 \*\*\*\*\*  
686 ++++++  
686 \*\*\*\*\*

Query:

Subject:

TRIPLE-HELICAL REGION 3 (COL3) .  
COL18A1  
Mature chain  
NONHELICAL REGION 4 (NC4) .

680 EKGARPGNSVGEKDPGNRGLPGPPGKKQAGPPGVMDPPPPPPGPPGCTMGLGF 739  
743 +KG+ G G KG G+ G G PGK G AG PG +GPPPPPPPPGPPG GF  
743 QKGSVDVGIPIPKSKGDLGPIGMPKSGLAGSPPVGPPGPPPPGPPG--FAAGF 800  
743 \*\*\*\*\*  
743 ++++++  
743 \*\*\*\*\*  
796 \*\*\*\*\*

Query:

Subject:

NONHELICAL REGION 4 (NC4) .  
COL18A1  
Mature chain  
TRIPLE-HELICAL REGION 4 (COL4) .

740 EDTGSGS--TQLNEPKLSRPTAIGLKKEG----- 770  
801 +D EGSG T + L P + GLKG+ G  
801 DMEGSGIPLWTTARSSDGLQGPSPGLKDPGAVAGLPAGKEVGADGAQGIIPGPPGRE 860  
801 \*\*\*\*\*  
801 ++++++  
801 \*\*\*\*\*  
819 \*\*\*\*\*

Query:

Subject:

TRIPLE-HELICAL REGION 4 (COL4) .  
COL18A1  
Mature chain  
NONHELICAL REGION 5 (NC5) .  
Conflict

771 ----DRGPKGERGM-----DGASIVGPPGPPGPHIKVLSNLSLINTHGFMNFS 817  
861 GPKGE+GM DG G PGP GPPG + +S S+  
861 GAAGSPGPKGKGMPEKGNPKKDVGRRPLPGPPGPPGVIVS-----SE 907  
861 \*\*\*\*\*  
861 ++++++  
861 \*\*\*\*\*  
902 \*\*\*\*\*  
900 \*

Query:

Subject:

Conflict  
NONHELICAL REGION 6 (NC6) .  
Conflict  
TRIPLE-HELICAL REGION 5 (COL5) .  
COL18A1  
Mature chain  
NONHELICAL REGION 5 (NC5) .

818 IPELVGPPGPDGLPGLPGFPGRGPKDGTGLPGFPLKGEQGEKEGEPGAILTEDIPL--E 875  
908 +V PGP+G PG GFP GP GPKGD G G GL G +GEKGE PG I + D  
908 DKAIIVSTPGPEGKPGVAGFPGPAGPKDLGSKGEQGLPGPKGEKEGEPGTIFSPDRALGH 967  
964 \*\*\*\*\*  
958 \*\*\*\*\*  
947 \*\*\*\*\*  
916 \*\*\*\*\*  
908 ++++++  
908 \*\*\*\*\*  
908 \*\*\*\*\*  
908 \*\*\*\*\*

Query:

Subject:

876 RLMGKGEPPGMHAGPMPGPKGPPGHKGEFGLPGRPPRGINGLKGTKGDPG----VIM 930  
968 G KGE PG G PGP G PGHKE G PGRPPRG NGKLG KG+PG M  
PQKAKGEFGRPPGPPYR--PGHKELIGFGRPPRGRTNGLKGEKEGEPGDASLFSM 1024

```
971      * * * * *
968      * * * * *
968      + + + + +
968      * *
```

```

931 QGPPGLPGPPGPSPGPAVINIKGAIFLIPVRPHCKMPVDTAHPGSPELLTFHGKGEKG 990
    + G PG PGPDPGPDPG I A V++ PG P G++G
1025 RGLPDPGPDPGPDPGPCMPLDYSNAF-----VESGRGPLP-----GGQG 1063
1044 *****
1025 *****
1025 *****
1025 ++++++++
1025 *****
1054 *****

```

```

991  SWGLPGSGKEGKEDGAGGPPGP-PLIDLAYLRHFLNNLNKENGNDKGFKEGKEKEDINGSF 1049
      G G KG+KG+ G GPPG P+DL +L + KG+ GD G KGE+GE G G F
1064  VQGSPGPKKDGEVGPGRPPGQFPIDL.FHLEAEMKGDKGDRGDAQGKGERGEPCAPGGGF 1123
1064  *****
1064  ++++++
1064  *****
1123  *****
1104  *****
1099  *****
1087  *****

```

```

1050 LMS-- --GPPGLPGNPAGQKGETVWGDPGPGPPAGPGLPBPFGFRPGDPGPPGPPGPP 1106
      S      GPPG PG PG G KGE++ +GPPG PG GPPG G G GPPGPPGPPG
1124 FSSSVBPBPBPBPBPGLBPCKGESI-- --RBPBPBPBPBPBPBPBPBPBPBPBPBPBPBP 1180
1124 *****
1124 ++++++
1124 ++++++
1124 ++++++
1130 ++++++
1130 ++++++
1157 ++++++

```

```

1107 PALLGA--AVALPGPPGPPGQPLPG---SRNLVTAFSNMDDMLÖKAHIVIEGTFIYL 1159
P+ G V++GGPPPG PG PG S V ++ ML K V EG I++
1181 PSFPGHRQTAVSPGPFGPPGPPGPMGASAGQVRWAIYQTMIDRKIREVPEGWLIFV 1240
1182 *****
1181 *****
1181 *****
1181 ++++++++
1181 * *****
1195 *****
1213 *****

```

```

1160 RDSTFEFFIRVRDGMKQLQELGELIP-----ADSP-1190
      + E ++RV+G++K+ L +P A PP
1241 AERRELVVRVNRNGFRKVLLEARTALPRGTGNEVALPLDPVLQHEGSPYTRREYSSTAR 1300
1266 *
1241 *****
1241 *****
1241 *****
1241 *****
1276 *****

```

with

Query:

Subject: NONHELIICAL REGION 11 (NC11).

COL18A1

Mature chain

Query:

Subject:

Mature chain NONHELIICAL REGION 11 (NC11).

COL18A1

Mature chain

disulfide

Query:

Subject:

disulfide

Conflict

disulfide NONHELIICAL REGION 11 (NC11)

Mature chain

COL18A1

Mature chain

Query:

Subject:

disulfide

disulfide NONHELIICAL REGION 11 (NC11).

Mature chain

COL18A1

Mature chain

Time: 0.82 user secs.

0.05 sys. secs.

0.87 total secs.

Gaped

Lambda

0.311 K H

0.422

Gaped

Lambda

0.267 K H

0.140

Matrix: BLOSUM62

Gap penalties: Existence: 11, Extension: 1

Number of Hits to DB: 657,831

Number of Sequences: 0

1191 -----PPALSSN-----PHQL-----LPPNPISSEAN-----KRALHLAALNM 1225  
PP L PH LPP P S + +P LHL ALN

1301 PWRADDILANPPRLPDQPPYPPGVPHHSSVYHLPPARPTLSLAHTHODFOQVHLVALNT 1360

1301 \*\*\*\*\*

1301 \*\*\*\*\*

1301 \*\*\*\*\*

1344 \*\*\*\*\*

1226 PFGSDIR---ADFOCFKORARAAGLSTYRAFLSSHLQDLSTIVRKARYSLPIVNLKGQ 1281

P SG +R ADFOCF+OARA GL T+RAFLSS LQDL +IVR+A+R S+PIVNLK +

1361 PLSGMRGIRGADFOCFQOARAVGLSGTPRAFLSSRLQDLYSIVRADRGSVPIVNLKDE 1420

1361 \*\*\*\*\*

1361 \*\*\*\*\*

1361 \*\*\*\*\*

1376 \*\*\*\*\*

1282 VLFNNWDSIFSGHGQFMNHIPIYSFDGRDITDPSWPQKVIWHGSSPHGVRLVDNYCEA 1341

VL +WDS+FSG GQ I+SPDGRD++ P+WPK +MHGS P G RL++YCE

1421 VLSPSWDSLFSGSQGLQPGARIFSFDGRDVLRLHPWPQKSVWHGSDPSGRRLMESYCE 1480

1478 \*\*\*\*\*

1437 \*\*\*\*\*

1421 \*\*\*\*\*

1421 \*\*\*\*\*

1421 \*\*\*\*\*

1421 \*\*\*\*\*

1342 WRTADTAVTGLASPLSTGKILDQKAYSCANRLIVLCIENSEFMTDARK 1368

WRT T TG AS L +G++L+OKA SC N IVLCIENSEFMT K

1481 WRTETGATGQASSLSLGRLLQKASCHNSYIVLCIENSEFMTSPK 1527

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

1481 \*\*\*\*\*

Number of extensions: 6996  
 Number of successful extensions: 1815  
 Number of sequences better than 10.0: 1  
 Number of HSP's better than 10.0 without gapping: 1  
 Number of HSP's successfully gapped in prelim test: 0  
 Number of HSP's that attempted gapping in prelim test: 0  
 Number of HSP's gapped (non-prelim): 1  
 length of query: 1388  
 length of database: 239,157,582  
 effective HSP length: 135  
 effective length of query: 1253  
 effective length of database: 215,896,677  
 effective search space: 270518536281  
 effective search space used: 270518536281  
 T: 9  
 A: 40  
 X1: 16 ( 7.2 bits)  
 129 (49.7 bits)  
 129 (49.7 bits)  
 S1: 42 (21.7 bits)  
 S2: 79 (35.0 bits)